Asymmetry in Subseasonal Surface Air Temperature Forecast Error with Respect to Soil Moisture Initialization

RANDAL D. KOSTER,[a] ANTHONY M. DEANGELIS,[b] SIEGFRIED D. SCHUBERT,[a,b] AND ANDREA M. MOLOD[a]

[a] *Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[b] *Science Systems and Applications, Lanham, Maryland*

(Manuscript received 5 February 2021, in final form 8 July 2021)

ABSTRACT: Soil moisture (*W*) helps control evapotranspiration (ET), and ET variations can in turn have a distinct impact on 2-m air temperature (T2M), given that increases in evaporative cooling encourage reduced temperatures. Soil moisture is accordingly linked to T2M, and realistic soil moisture initialization has, in previous studies, been shown to improve the skill of subseasonal T2M forecasts. The relationship between soil moisture and evapotranspiration, however, is distinctly nonlinear, with ET tending to increase with soil moisture in drier conditions and to be insensitive to soil moisture variations in wetter conditions. Here, through an extensive analysis of subseasonal forecasts produced with a state-of-the-art seasonal forecast system, this nonlinearity is shown to imprint itself on T2M forecast error in the conterminous United States in two unique ways: (i) the T2M forecast bias (relative to independent observations) induced by a negative precipitation bias tends to be larger for dry initializations, and (ii) on average, the unbiased root-mean-square error (ubRMSE) tends to be larger for dry initializations. Such findings can aid in the identification of forecasts of opportunity; taken a step further, they suggest a pathway for improving bias correction and uncertainty estimation in subseasonal T2M forecasts by conditioning each on initial soil moisture state.

SIGNIFICANCE STATEMENT: Not all forecasts are created equal. Even before a given forecast is produced, the nature of its initial conditions may indicate that it will prove more accurate than corresponding forecasts started at other times. We address here how the character of the soil moisture at the beginning of a forecast may provide such information. We find that under certain conditions, when the initial state of the soil is wet, the bias in the forecast is reduced and, to a lesser extent, the random error in the forecast is also reduced. Knowing ahead of time when to put more trust into a forecast should be of substantial benefit to forecast end-users.

KEYWORDS: Hydrometeorology; Surface temperature; Forecast verification/skill; Forecasting; Land surface model

1. Introduction

a. Background and problem statement

The ocean–atmosphere–land system is inherently chaotic (Shukla 1998). Nevertheless, its current state contains information about how it may evolve, and this information, if properly utilized, can provide for skillful Earth system forecasts. Characterizing the evolution of the Earth system from an initial state—characterizing the inherent predictability of Earth system variables at multiple time scales in the face of chaos—continues to be a profound scientific challenge. Naturally, society could benefit from such a characterization, e.g., through the improved quantification, and perhaps reduction, of meteorological forecast uncertainty.

Traditional weather forecasts, extending out to about 10 days, extract most of their skill from the initialization of the atmosphere (Leutbecher and Palmer 2008; Zhang et al. 2019). In contrast, forecasts at seasonal leads (>2 months) extract their skill from slowly changing modes in the coupled atmosphere–ocean system (Shukla 1998). Nestled between the weather forecast and the seasonal forecast is the subseasonal forecast, with leads of 2 weeks–2 months. Subseasonal forecasts, a relatively recent research focus (Robertson et al. 2015; Vitart et al. 2017; Pegion et al. 2019), rely on some unique sources of skill, including the development and evolution of the Madden–Julian oscillation (Kim et al. 2018) and stratospheric sudden warming events (Scaife et al. 2016). Another important source is the persistence of soil moisture anomalies (Koster and Suarez 2001; Seneviratne et al. 2006), given that a persisted anomaly may be able to influence meteorological conditions through land–atmosphere coupling (Koster and Suarez 2003; Dirmeyer and Halder 2017). For example, under certain conditions, an anomalously wet soil may produce an anomalously high evapotranspiration (ET) rate that helps cool the land surface (Seneviratne et al. 2010), leading to an anomalously low near-surface air temperature (T2M).

It is now generally accepted that useful forecast skill at subseasonal time scales is very likely episodic, so that progress in improving skill will rely on our ability to identify, in advance, forecasts of opportunity—those times during which one or more of the various physical phenomena discussed above allow for increased predictability (Mariotti et al. 2020). With this in mind, we focus here on the connection between initial soil moisture conditions and T2M forecasts at subseasonal lead times, taken here to be days 16–30 of a forecast. While the contribution of soil moisture initialization to both T2M and precipitation forecast skill has been examined extensively in the literature (e.g., Koster et al. 2011; van den Hurk et al. 2012;

Corresponding author: Randal D. Koster, randal.d.koster@nasa.gov

DOI: 10.1175/JHM-D-21-0022.1

Hirsch et al. 2014), here we examine a heretofore unexplored facet of the problem: the degree to which the character (i.e., wet versus dry) of the initial soil moisture conditions imprints itself on this skill. We aim to determine if forecast skill is indeed higher for initial states of a certain character; such initial states could then arguably serve as markers for forecasts of opportunity.

We propose herein two distinct mechanisms by which soil moisture might affect forecast skill levels, mechanisms that affect different components of the total forecast error. Total forecast error can be quantified using the root-mean-square error (RMSE) metric relative to observations. The RMSE can be separated (Entekhabi et al. 2010a) into our two components of interest:

$$\mathrm{RMSE}^2 = \mathrm{bias}^2 + \mathrm{ubRMSE}^2, \quad (1)$$

where bias is the average difference between the (uncalibrated) forecast and the observations and ubRMSE is the unbiased root-mean-square error, a measure of the forecast system's random error. One of the mechanisms we propose acts directly and solely on the bias; indeed, through this mechanism, the model's forecast bias can be interpreted as conditional on the initial soil moisture state (see section 2e). The other mechanism acts solely on the ubRMSE.

The overall goal of this study can be stated as follows. Having described the two mechanisms, we will search for signatures of both in a suite of T2M forecasts produced with a state-of-the-art subseasonal-to-seasonal (S2S) forecast system. The signatures, if found, will support the idea that relatively dry and relatively wet soil moisture initial states lead to distinct T2M forecast error behavior—that is, that the initial soil moisture state provides information about the accuracy of a given forecast.

b. *Hypothesized mechanisms*

Our two proposed mechanisms involve the canonical relationship between soil moisture and ET efficiency, shown in Fig. 1. ET efficiency is the fraction of the land surface's net radiative energy used to evaporate moisture from the surface; that is, evapotranspiration efficiency ε is $\lambda\mathrm{ET}/R_{\mathrm{net}}$, where λ is the latent heat of vaporization and/or fusion and R_{net} is the net radiation. The salient feature of the canonical relationship is the presence of two distinct regimes: a soil moisture–controlled regime (the "dry regime"), wherein ET efficiency increases with soil moisture, and an energy-controlled regime (the "wet regime"), wherein soil moisture is no longer the bottleneck to ET, so that ET efficiency is largely insensitive to soil moisture variations. The nature of the relationship has been studied extensively in the literature (e.g., Eagleson 1978; Salvucci 2001). While it is not (usually) explicitly coded into land surface models, it is nevertheless implicitly captured by a model's numerous interacting parameterizations (Dirmeyer et al. 2006; Koster et al. 2019).

Figure 2 illustrates our first proposed mechanism, which focuses on T2M forecast bias. The top panel shows what might happen when the soil moisture is initialized wet (case 1). An unbiased (in terms of precipitation) forecast system might produce, within a single forecast, an ensemble of forecasted soil moistures with an average of W_{ub}, associated with an average forecasted evapotranspiration efficiency of $\varepsilon_{\mathrm{ub}}$. The actual forecast system, however, might have a negative precipitation bias that manifests itself at subseasonal leads, and the resulting forecasts of W and ET efficiency (W_b and ε_b in the panel) would accordingly be lower. The key point to notice is that the nonlinearity in the relationship induces only a small difference between $\varepsilon_{\mathrm{ub}}$ and ε_b, despite a large difference between W_{ub} and W_b. That is, for a wet soil moisture initial state, the negative precipitation bias does not lead to a large bias in forecasted ET efficiency.

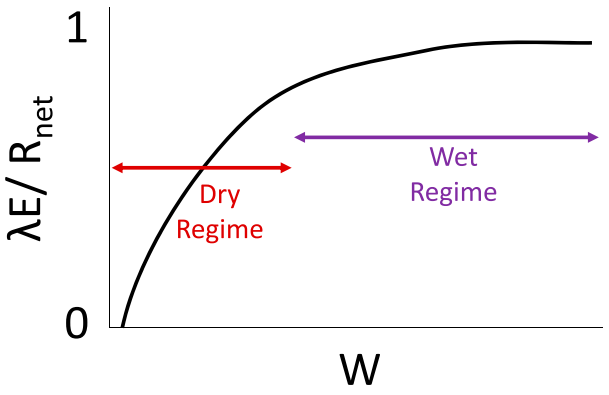

FIG. 1. Canonical relationship between soil moisture W and the evapotranspiration efficiency $\lambda\mathrm{ET}/R_{\mathrm{net}}$. Soil moisture variations have a much stronger impact on ET variations in the dry regime.

Consider now the contrasting impact of a dry soil moisture initial state (case 2, lower panel)—when initialized dry, the forecast system, faced with the same precipitation bias at subseasonal leads, would produce a large bias in the ET efficiency. The stark distinction in ET efficiency errors between cases 1 and 2 is, of course, a direct result of the nonlinearity of the relationship in Fig. 1. To the extent that the net radiation itself shows relatively low variability when averaged over days 16–30 of a forecast (usually a reasonable assumption, but see further discussion in section 2c), we see that wet and dry soil moisture initial states should have a distinctly different impact on the ET bias itself.

What should we accordingly expect regarding bias in forecasted air temperature? Again, the connections between ET and T2M are well established (Seneviratne et al. 2010); in the soil moisture–controlled ET regime (Fig. 1), lower soil moistures lead to lower ET rates that in turn induce higher T2M values through a decrease in evaporative cooling. Because the negative precipitation bias induces a significant negative ET bias for case 2, we should expect for case 2 a concomitant positive increase in the T2M bias. For case 1, on the other hand, the precipitation bias should have little effect on the T2M bias, at least through the ET pathway.

Figure 3 illustrates the second proposed mechanism, the one that affects ubRMSE. We cannot, of course, know at the start of a forecast what the true soil moisture will be at the subseasonal lead; due to the chaotic nature of the coupled Earth system, the true soil moisture may take on a range of values, represented in the top panel of Fig. 3 by the red rectangle on the x axis. (Note that in an unbiased forecast system, this

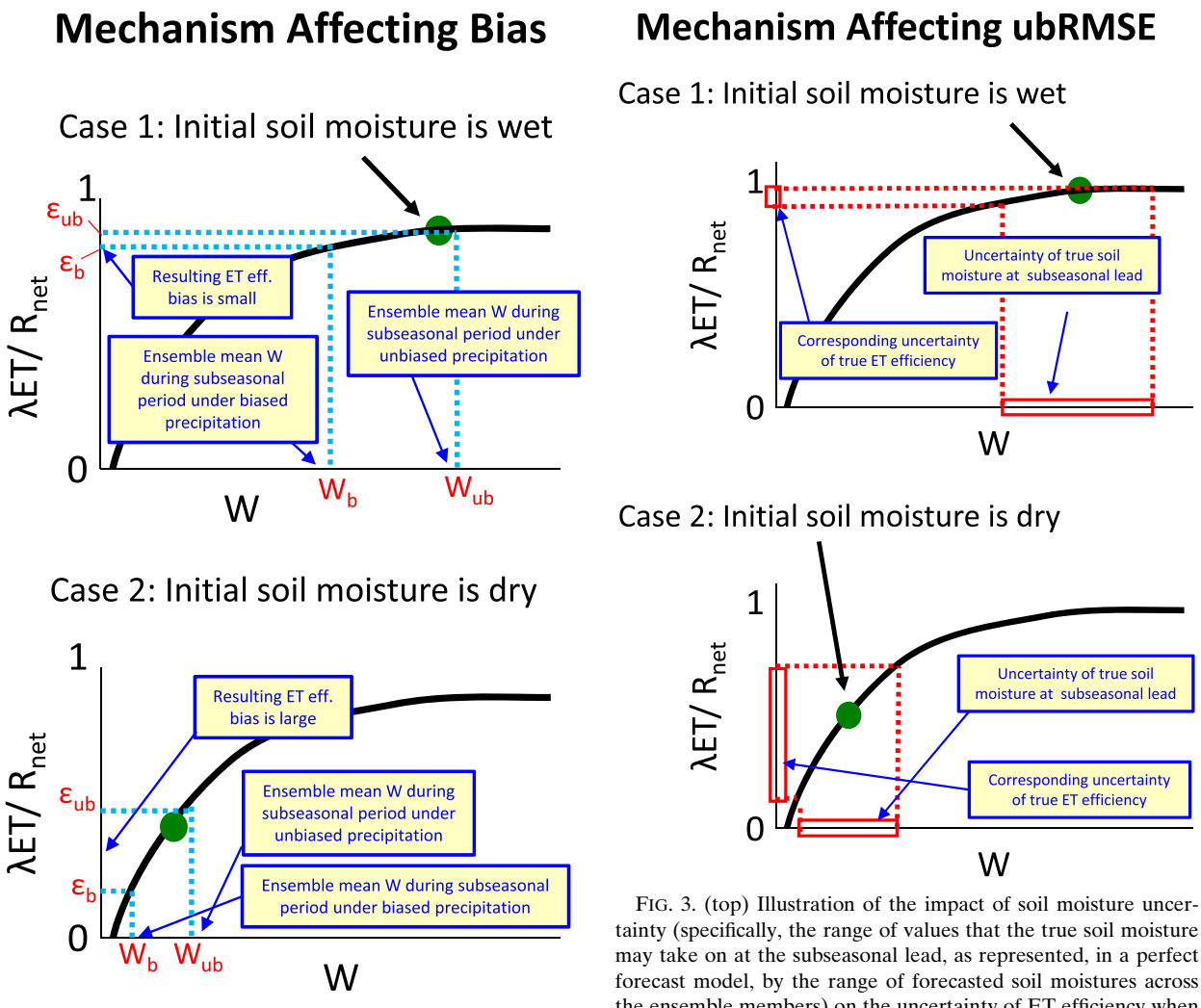

FIG. 2. (top) Illustration of the impact of a precipitation deficit on ET efficiency when the soil moisture for a specific forecast begins in the wet regime. The terms W_{ub} and ε_{ub} represent ensemble mean soil moistures and ET efficiencies, respectively, for that specific forecast under unbiased precipitation; W_b and ε_b are the corresponding values under biased precipitation. (bottom) As in the top panel, but for when the soil moisture begins in the dry regime.

FIG. 3. (top) Illustration of the impact of soil moisture uncertainty (specifically, the range of values that the true soil moisture may take on at the subseasonal lead, as represented, in a perfect forecast model, by the range of forecasted soil moistures across the ensemble members) on the uncertainty of ET efficiency when the soil moisture begins in the wet regime. (bottom) As in the top panel, but for when the soil moisture begins in the dry regime. The larger uncertainty of ET efficiency in the bottom panel is indicative of a higher ubRMSE for ET efficiency and thus, by extension, a higher ubRMSE for T2M.

uncertainty range is characterized by the intraensemble spread in the ensemble forecast.) The top panel (case 1) represents the case for which the expected value of soil moisture is in the wet regime; as seen in the plot, for this case, the large uncertainty range for soil moisture gets translated to only a narrow uncertainty range for the ET efficiency. In contrast, the bottom panel shows the case of a dry expected soil moisture value (case 2); here, the corresponding uncertainty range for the ET efficiency is large. As we saw above for ET bias, the difference in the uncertainty ranges for ET efficiency (and thus for ET) between cases 1 and 2 stems entirely from the nonlinear shape of the canonical ET efficiency–soil moisture relationship.

Because the examples in Fig. 3 focus on an idealized, unbiased system, the distinction between cases 1 and 2 in the uncertainty ranges (ensemble spreads) for ET translates to a distinction in the ubRMSE of forecast ET. This is because again, the ensemble spread represents the range of possible ET values that the single realization of nature might take on. A larger spread suggests a greater likelihood that nature's realization of ET will lie farther away from the forecast ensemble mean. A larger spread accordingly suggests a larger ubRMSE for the forecasted ET.

Considering now the ability of ET anomalies to translate themselves into T2M anomalies, Fig. 3 suggests that a dry soil moisture (and thus a dry soil moisture initial state in conjunction with soil moisture memory) could potentially lead to greater ensemble spread in the T2M forecast and thus, via the reasoning outlined above, to a larger forecast T2M ubRMSE. [Consistent with this idea, Koster et al. (2020) document a larger intraensemble spread for forecast T2M

under dry conditions.] Again, we must emphasize that Fig. 3 is highly idealized; the actual T2M ubRMSE will also be affected by other chaotic elements of the system and by various model deficiencies. Still, these additional factors would presumably act similarly for dry and wet soils. Even in the presence of these other factors, the mechanism outlined in Fig. 3 should induce a distinction in T2M ubRMSE based on initial soil moisture content.

Together, Figs. 2 and 3 capture the behavior we will seek in the subseasonal forecasts produced with the state-of-the-art forecast system: for relatively dry soil moisture initial states, we expect to see a higher ubRMSE for T2M forecasts and, for cases with a precipitation bias, a higher bias for T2M forecasts. Given the additive contributions of (squared) bias and ubRMSE to the total RMSE, we thus also expect to see higher total error in forecast T2M under drier soil moisture initial states.

c. *Structure of paper*

This paper is organized as follows. Section 2 below describes the forecast system used in our analysis, the observational data used for skill assessments, and the approach used to distinguish soil moisture initial conditions at a given location as ''dry'' or ''wet.'' It also shows where on the globe the mechanisms above can be reasonably evaluated. Section 3 presents the results, focusing on the impacts of dry versus wet initial states on T2M forecast bias and forecast ubRMSE separately. Additional discussion is provided in section 4, and conclusions are provided in section 5.

2. Data and methods

a. *GMAO systems and products*

The Global Modeling and Assimilation Office (GMAO) at the National Aeronautics and Space Administration's Goddard Space Flight Center (NASA/GSFC) hosts two products of relevance to this work: (i) a state-of-the-art reanalysis known as the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al. 2017), and (ii) a series of subseasonal forecasts performed with the GMAO's seasonal forecast system (Molod et al. 2020).

MERRA-2 assimilates a wide variety of conventional and satellite-based measurements (McCarty et al. 2016), including aerosol information (Randles et al. 2016), into a global atmosphere/land surface modeling environment [the Global Earth Observing System, or GEOS, model (Molod et al. 2015)] to produce a comprehensive and consistent picture of atmospheric and land surface fields over the period from 1980 to present. The present analysis employs the following daily-averaged MERRA-2 products on a 0.5° latitude × 0.625° longitude grid: evapotranspiration [ET, from GMAO (2015a)], average profile soil moisture [W, from GMAO (2015a)], gauge-corrected precipitation [P, from GMAO (2015b)], and net radiation [R_{net}, from GMAO (2015c)]. The ET, W, and R_{net} fields are used to characterize dry versus wet soil moisture regimes at each grid cell (section 2b). While in some ways ET is more directly controlled by root zone moisture, the profile moisture W is a valid surrogate when constructing ET–soil moisture relationships for the purpose of identifying dry versus wet regimes; this is because root zone moisture and profile moisture covary strongly in this model (Koster et al. 2020) and because W provides a better indication of the total land moisture present at the start of a forecast. MERRA-2 P data, used to quantify precipitation bias, are considered a suitable representation of truth, given that MERRA-2 explicitly includes corrections to P from rain gauge networks (Reichle et al. 2017) and that we focus here only on areas for which rain gauge density is adequate (section 2c). Note that these P corrections are implicitly built into MERRA-2's ET and W fields.

GMAO's S2S version 2 prediction system (Molod et al. 2020) initializes a fully coupled ocean–atmosphere–land–sea ice model using a weakly coupled Atmosphere–Ocean Data Assimilation System. During the analysis that provides the initial states for the forecasts, the precipitation (as in MERRA-2) is scaled to agree with a rain gauge product (Reichle et al. 2017), adding realism to the initial soil moisture anomalies used in the forecasts. The land model component of the forecast system [the Catchment model of Koster et al. (2000a)] is the same as that used for MERRA-2, so the characterization at each grid cell of wet versus dry soil moisture regimes derived from MERRA-2 data (section 2b) also applies to the S2S system. The Catchment model features an explicit treatment of subgrid soil moisture variability; the land surface is characterized by up to three hydrological regimes whose areas vary in time according to the overall moisture state and the existing topography: a saturated regime (from which evapotranspiration, ET, occurs without soil moisture stress), a subsaturated regime (from which ET is limited by soil moisture stress), and a wilting regime (from which transpiration is prevented). Photosynthesis physics are not directly incorporated into this version of the model, being instead captured by empirical formulations. A full, implicit surface energy balance calculation ties ET to the land surface temperature (and accordingly to the overlying air temperature). Higher ET rates imply that more of the available radiative energy is used for the phase change of water, leaving less available to heat the surface; as a result, higher ET rates induce cooler temperatures.

Again, we focus in this study on the series of boreal warm season forecasts (actually, hindcasts) produced by the GEOS S2S system. For each warm season month (June–August), we analyze a total of 102 independent forecasts: six forecasts in a given year (with start dates spaced 5 days apart during the month in question) for each of the 17 years in the period 1999–2015. The 17-yr period considered corresponds to the analysis period covered by the SubX multimodel forecasting experiment (Pegion et al. 2019). While each of the 102 forecasts considered for a given month consists of 4 ensemble members, the method used to perturb the initial conditions for the ensemble members does not include a perturbation of W, so we distinguish 102 initial W states per month. We focus here on averages over forecast days 16–30, i.e., forecasts at our chosen subseasonal lead. All model simulations were performed on a cubed sphere grid, approximating a 0.5° × 0.5° resolution;

model output data were regridded onto the 0.625° × 0.5° grid used for MERRA-2 diagnostic products.

b. Categorization of dry and wet forecast initial states

While the nonlinear relationship in Fig. 1 is not explicitly coded into the Catchment land surface model of the GEOS forecast system, evidence that it is nevertheless in operation can be derived through analysis of the model's ET, R_{net}, and soil moisture output diagnostics (Koster et al. 2019, 2020). For the present analysis, it is critical to estimate from the model diagnostics the particular soil moisture $W_{boundary}$ that, for a given grid cell and time of year, divides the full soil moisture range into Fig. 1's dry regime and wet regime. Only by identifying this boundary can a given forecast's initial soil moisture be categorized as "dry" or "wet."

We derive $W_{boundary}$ through analysis of MERRA-2 ET, R_{net}, and W values; as noted above, this is reasonable given that MERRA-2 uses the same land surface model as does the GEOS S2S system. To help ensure independence from the forecast suite, $W_{boundary}$ is derived here from MERRA-2 data lying outside the forecast period (i.e., MERRA-2 data from 1980 to 1998). Highlights of the approach we used as applied at a representative grid cell (in the central United States) and representative month (June) are illustrated in Fig. 4. We first plot daily values of W against corresponding values of $\lambda ET/R_{net}$ from all June days in 1980–98. The wettest 10% of soil moistures in the plot are then identified, and the mean soil moisture and ET efficiency for that subset are used to position the point A in the figure. Similarly, averages over the driest 10% of soil moistures are used to position the point B. To determine $W_{boundary}$, 20 evenly spaced locations between the points A and C are tested one by one; we identify the location (point D in the figure) such that, when the piecewise linear function (in dark blue) is fitted through A, D, and B, the function provides the closest approximation to the scatter of points in the plot. In essence, we determine the point D on the line AC such that the RMSE between the piecewise function and the individual $\lambda ET/R_{net}$ values is minimized. Note that if the slope of the line DB is too shallow (i.e., less than 1), we assume that the two regimes are not properly distinguished and do not consider the grid cell in our analysis for the month in question.

This, of course, is but one potential approach for identifying the proper value of $W_{boundary}$; other approaches are also available (e.g., Akbar et al. 2018). We tested the sensitivity of our study's results to the precise value of $W_{boundary}$ by modifying it by 0.02 in either direction and then repeating our analyses. We find that while the modification of $W_{boundary}$ results in small quantitative differences in the fields we generate (not shown), the conclusions we extract from the fields remain unchanged, providing confidence that the approach is acceptable for this study. We note also that although $W_{boundary}$ varies substantially across the globe, we find (not shown) no clear predilection for the mechanisms in Figs. 2 and 3 to operate most effectively within a specific range of $W_{boundary}$ values.

c. Focused study area

It is important at the outset to identify the locations at which we can reliably distinguish the impacts of wet and dry soil moisture initial states on forecast skill. Indeed, for an evaluation of the proposed mechanisms to be viable at a given location, four features must be in place. First, the net radiation must be large enough to translate the W differences into sizable ET differences for a dry initial condition. Second, across the forecast suite, there must be a number of initial W values in the dry regime as well as a number in the wet regime to allow a robust averaging of results for each regime. Third, in the dry regime, ET must tend to increase as expected with soil moisture. Fourth, the rain gauge density at the location must be adequate. This last condition relates to the fact that the GMAO S2S forecast system does not use satellite-based precipitation estimates or soil moisture retrievals in the simulations providing the initial W values applied in the forecasts; in this system, the accuracy of the initial soil moistures relies instead on precipitation as measured by gauges. An adequate rain gauge density is needed for our analysis because there is no point evaluating how soil moisture affects T2M forecast skill if the accuracy of the initial soil moisture is itself in question.

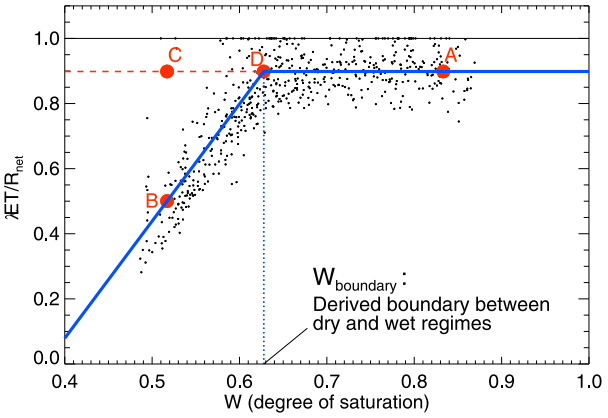

FIG. 4. Illustration of the calculation of $W_{boundary}$, the soil moisture (degree of saturation) delimiting the dry and wet regimes (see text).

We address the net radiation requirement by focusing on the boreal warm season of June through August. [Evidence for the operation of the mechanisms in May and September, the shoulder months, is present but weaker, as expected (not shown).] Locations satisfying the other three requirements are illustrated in Fig. 5. First, Fig. 5a shows the locations (in red) where at least 5 of the 102 June forecast start dates (6 June start dates per year for all 17 SubX forecast years) have initial W values in the dry regime and at least 5 have initial W values in the wet regime, as determined by comparing each of the initial moisture states to the computed $W_{boundary}$ values. Requiring more than 5 start dates, of course, would allow for more robust statistics, but it would also reduce the size of the area that could be considered; through analysis of maps generated with different requirements, we settled on a minimum of 5 start dates as an optimal compromise. Note that while much of North America satisfies this criterion, a smaller fraction of the grid cells in Eurasia do; for this modeling system, many desert and high latitude locations in Eurasia could not be included in our analysis.

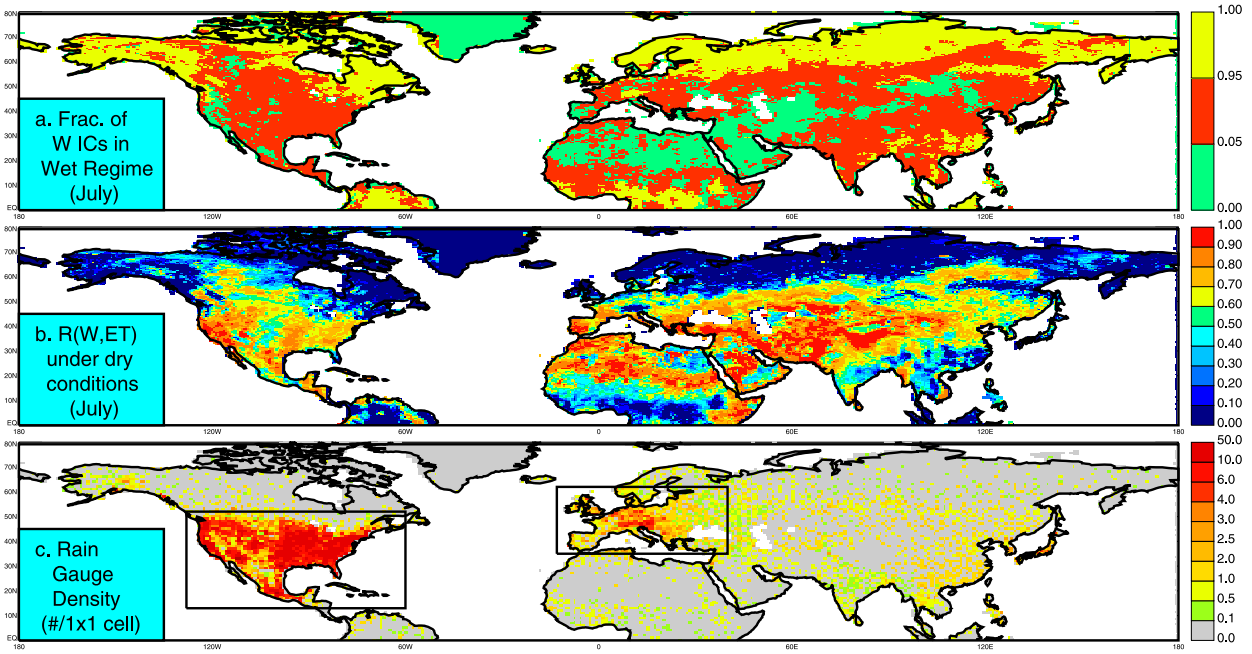

FIG. 5. (a) Fraction of the 102 July forecasts for which the initialized soil moisture lies in the wet regime (see Fig. 1). (b) Correlation between daily profile soil moisture and daily ET for the subset of dates with soil moisture below W_{boundary}, from MERRA-2 data holdings for July during 1980–98. (c) Rain gauge density underlying the precipitation data used to initialize the subseasonal forecasts (number of gauges per 1° × 1° cell, regridded here to 0.625° × 0.5°).

Figure 5b addresses the question of whether ET does tend to increase with soil moisture in the dry regime at a given location. (A reasonable connection between the two variables is required, of course, for the mechanisms in Figs. 2 and 3 to operate.) Specifically, Fig. 5b shows – for those July days during 1980 through 1998 in which the soil moisture at a given location is below the W_{boundary} value determined for July (section 2b)—the correlation between daily soil moisture and daily ET. This test is necessary because a high correlation is far from guaranteed; again, the canonical relationship in Fig. 1 is not explicitly coded into the land model. Furthermore, the relationship in Fig. 1 is not between soil moisture and ET but rather between soil moisture and $\lambda\text{ET}/R_{\text{net}}$, so that if the net radiation itself varies excessively from day to day, ET might not increase with increasing soil moisture as needed. Despite these potential issues, the correlations in Fig. 5b are, as expected, positive almost everywhere and above 0.7 in most locations. The values, however, are low in high latitude regions and in the tropics, as well as in (for example) central Europe. The low values that appear in many of these areas presumably stem from a limited number of dry cases over which to compute the correlation, as suggested in Fig. 5a.

Finally, Fig. 5c shows the density of the rain gauges underlying the initialization of the GEOS subseasonal forecasts—the rain gauges that feed into the 0.5° × 0.5° daily Climate Prediction Center Unified (CPCU) precipitation product (Xie et al. 2007; Chen et al. 2008). While the density is high in the conterminous United States (CONUS) and across Europe, and while it is reasonable in some other regions (e.g., India, eastern China), the density is low across much of the globe, a long-recognized problem (Kidd et al. 2017). Note, however, that the actual use of these gauge data in GEOS systems is a little complex, as described in detail by Reichle et al. (2017). In short, for the initialization process (and indeed for the production of the MERRA-2 precipitation rates used for our calculation of precipitation biases), the gauge data are not used north of 62.5°N; here, GEOS analysis data (Lucchesi 2018) are used instead. Between 42.5° and 62.5°N, a tapered blend of the analysis and gauge data is used. Also, in Africa, CPC Merged Analysis of Precipitation–corrected data (Xie and Arkin 1997) rather than CPCU-corrected data are used.

For clarity of presentation, the fields generated in our analysis will focus on the boxed area over CONUS in Fig. 5c, an area consisting of grid cells that, for the most part, satisfy all of the above criteria. Because parts of Europe (particularly the Balkan Peninsula) are also potentially relevant, the boxed area over Europe will be examined briefly in section 4. In our plots, we will mask out grid cells colored yellow or green in Fig. 5a.

d. *Verification data*

We compare forecasted T2M at subseasonal leads to the independent, fully observations-based CPC temperature dataset (https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globaltemp.html). The CPC data are provided at 0.5° × 0.5° resolution and are generated from station observations. We approximate a given day's observed average temperature using 0.5 (T_{min} + T_{max}), given that only the daily minimum temperature (T_{min}) and maximum temperature (T_{max}) are provided in the CPC dataset.

e. Calculation of error metrics

The determination of W_{boundary} allows us to divide the N forecasts of a forecast suite into two groups: those that are initialized in the wet regime (N_{wet}) and those that are initialized in the dry regime (N_{dry}). (As noted above, $N = 102$ for a given month.) We can accordingly derive a subseasonal forecast bias for each subset separately:

$$\text{bias}_{\text{wet}} = 1/N_{\text{wet}} \sum_{i=1,N_{\text{wet}}} (\text{T2M}_{\text{forecast},i} - \text{T2M}_{\text{observed},i}), \quad (2)$$

$$\text{bias}_{\text{dry}} = 1/N_{\text{dry}} \sum_{i=1,N_{\text{dry}}} (\text{T2M}_{\text{forecast},i} - \text{T2M}_{\text{observed},i}), \quad (3)$$

where $\text{T2M}_{\text{forecast},i}$ is the four-member ensemble mean forecast of T2M averaged over days 16–30 for forecast i, and $\text{T2M}_{\text{observed},i}$ is the corresponding observed temperature during that forecast period. With the biases for these subsets computed, we can quantify the impact of soil moisture initialization regime on bias using the difference $\text{bias}_{\text{dry}} - \text{bias}_{\text{wet}}$.

The overall precipitation and T2M bias maps to be shown later are based on equations similar to (2) and (3), but using all forecasts—not just those with wet or dry initial conditions. With a forecast period spanning 17 years, with 6 forecast start dates per month, and with 4 ensemble members per forecast, a total of 408 simulations with the forecast system at the subseasonal lead of interest are analyzed to produce the overall precipitation and T2M bias maps. As discussed above, the observed precipitation rates used for the bias calculation are the rain gauge-corrected rates included in the MERRA-2 product.

The RMSE values for the dry and wet initial states are computed with

$$\text{RMSE}_{\text{wet}} = \left[1/N_{\text{wet}} \sum_{i=1,N_{\text{wet}}} (\text{T2M}_{\text{forecast},i} - \text{T2M}_{\text{observed},i})^2 \right]^{0.5}, \quad (4)$$

$$\text{RMSE}_{\text{dry}} = \left[1/N_{\text{dry}} \sum_{i=1,N_{\text{dry}}} (\text{T2M}_{\text{forecast},i} - \text{T2M}_{\text{observed},i})^2 \right]^{0.5}. \quad (5)$$

Application of (1) then provides our equations for ubRMSE:

$$\text{ubRMSE}_{\text{wet}} = (\text{RMSE}_{\text{wet}}^2 - \text{bias}_{\text{wet}}^2)^{0.5}, \quad (6)$$

$$\text{ubRMSE}_{\text{dry}} = (\text{RMSE}_{\text{dry}}^2 - \text{bias}_{\text{dry}}^2)^{0.5}. \quad (7)$$

As with the bias metric, the differences $\text{ubRMSE}_{\text{dry}} - \text{ubRMSE}_{\text{wet}}$ and $\text{RMSE}_{\text{dry}} - \text{RMSE}_{\text{wet}}$ in the figures below quantify the impact of a dry initial state on the ubRMSE and RMSE, respectively.

We use a bootstrapping procedure to evaluate the significance of the computed differences. For a given grid cell and month, we randomly sample (without replacement) two sets of forecast and observed 15-day T2M (or precipitation) pairings from the 102 available pairings: one set numbering N_{dry} pairings, and the other numbering $102 - N_{\text{dry}}$, or N_{wet}, pairings. (Given the random sampling, the elements in neither set are tied to the initial soil moisture state.) We then compute the bias, ubRMSE, or RMSE differences for that random sampling using the equations above. The process is repeated 500 times to provide a distribution for the differences, a distribution that can be used to evaluate the statistical significance of the actual differences, i.e., those that are in fact keyed to the initial soil moisture state.

The approach used to make the random selections, by the way, also allows the determination of significance levels for the large-scale means of the differences. A given time period is associated with a spatial distribution of initial soil moisture as well as a spatial distribution of 15-day T2M pairings (forecast versus observed). The random selection process involves, in essence, a shuffling of the dates associated with the 15-day pairings while retaining the spatial information in the initial soil moisture fields and thus the information on which grid cells during a specific period are considered dry and which are considered wet. The upshot is that we account explicitly for the impact of relevant spatial correlations (in both the initial soil moisture fields and the 15-day meteorological fields) on the large-scale averages of the metrics.

3. Results over CONUS

We now evaluate the mechanisms illustrated in Figs. 2 and 3 using the suite of subseasonal forecasts produced by the state-of-the-art NASA GMAO GEOS S2S-2 forecast system (section 2a).

a. Forecast bias

Section 2c above identified CONUS as a region over which we might find evidence of the mechanisms outlined in Figs. 2 and 3. For the mechanism affecting the bias component of error (Fig. 2), CONUS is particularly worthy of study, as it features a particularly large negative precipitation bias during summer. A negative summertime precipitation bias in CONUS is, in fact, endemic to most modeling systems (Lin et al. 2017).

The first and second columns in Fig. 6 show, for the CONUS region, composite maps of subseasonal T2M forecast bias for forecasts initialized in the dry regime and the wet regime, respectively, at a given cell (using the method described in section 2b). Biases are computed relative to independent CPC air temperature observations (as described in sections 2d and 2e). Notice that for the majority of CONUS, the biases are clearly smaller for the wet initial states. That is, over CONUS, wet initial states provide T2M forecasts that are closer, in terms of mean bias, to the observations.

The third column in Fig. 6 shows the differences between the fields in the first and second columns, and the fourth column shows the overall precipitation bias produced by the forecast system at the 16–30 day subseasonal lead (section 2e). The mechanism outlined in Fig. 2 suggests that the T2M bias difference (third column) should be positive where the forecasted precipitation is, on average, biased negative (fourth column). The patterns do not match exactly (the pattern correlations range from only 0.33 in June to 0.49 in July for points north of 30°N [based on ~2000 points], and they range from 0.42 to 0.63 when considering points north of 30°N and east of 100°W

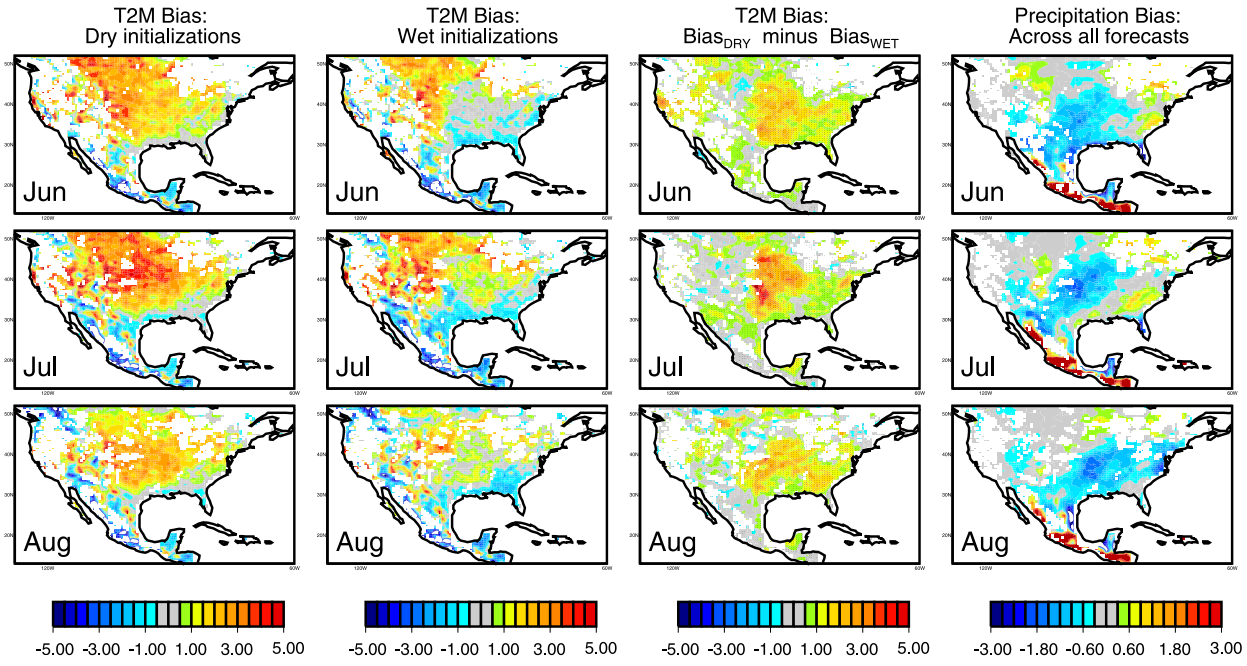

FIG. 6. (column 1) Average T2M forecast bias (K) generated by those forecasts with initial soil moisture in the dry regime (see Fig. 1). These are composite maps; the particular forecasts contributing to the results shown are grid cell specific. (column 2) As in column 1, but for those forecasts with initial soil moisture in the wet regime. (column 3) Differences: column 1 minus column 2. Dots indicate differences that are statistically significant at the 95% confidence level. (column 4) Precipitation bias (mm day^{-1}) at the subseasonal lead averaged across all forecasts.

[based on ~1200 points]), but this is to be expected given sampling constraints and the fact that local ET is not the only determinant of T2M across the continent. T2M is indeed affected by a variety of dynamical and thermodynamic factors, including potential impacts of remote ET anomalies (Koster et al. 2016). The salient result from the figure is rather the first-order agreement in pattern between the third and fourth columns—according to the investigated mechanism, the large negative precipitation deficit in the center of CONUS should lead to a large T2M bias difference there, and this is precisely what is seen. The T2M biases produced by the two forecast subsets, one initialized dry and the other initialized wet, differ by as much as 2–3 K, particularly in July, and the differences over about half the area shown are significant at the 95% level. The first-order agreement between the bias difference and precipitation deficit patterns, including their monthly variations, supports the idea that the mechanism outlined in Fig. 2 is indeed in operation.

To what extent, though, might the forecasted precipitation bias itself be affected by the soil moisture initialization, and how might this affect the T2M bias? Past analyses have shown that while soil moisture initialization has a considerably stronger impact on T2M forecasts than on precipitation forecasts at subseasonal leads, the impact on precipitation is nevertheless nonzero (Koster et al. 2011). If a dry soil moisture initial state does manage to induce a reduced amount of precipitation (and thus a more negative precipitation bias), the mechanism outlined in Fig. 2 would be enhanced, and the corresponding upward bump in the T2M bias would be increased. We investigated this idea by examining the precipitation biases over CONUS for the dry and wet soil moisture initial states separately. These differences, plotted in Fig. 7, indicate that overall, the impact of land surface feedback on precipitation, if present, is at most of secondary importance, though the small positive feedback that does appear to occur in central CONUS during July (statistically significant at the 95% level) is consistent with the higher T2M bias differences there for that month (relative to the other months) in the third column of Fig. 6. The convective precipitation scheme used in the GMAO S2S system, by the way, is the Relaxed Arakawa–Schubert scheme (Moorthi and Suarez 1992), a scheme that in past modeling studies has shown a strong propensity for land–atmosphere feedback (Koster et al. 2000b), to a degree that some would argue is excessive (Ruiz-Barradas and Nigam 2006). Based on our experience with this scheme, it seems unlikely that the results in Fig. 7 significantly underestimate the importance of precipitation feedback for this problem.

Whether or not feedbacks with precipitation are relevant, the strong summertime dry precipitation bias in central CONUS does allow the impacts of the mechanism outlined in Fig. 2 to be clearly visible—Fig. 6 supports the idea that the mechanism is indeed operating. We turn now to the mechanism outlined in Fig. 3, the one that addresses soil moisture impacts on the unbiased RMSE, or ubRMSE.

b. *Unbiased RMSE*

Figure 8 shows, for June, July, and August separately, the ubRMSE computed for dry initial states (column 1), wet initial states (column 2), and their differences (column 1 minus column 2). Of particular interest are the difference plots.

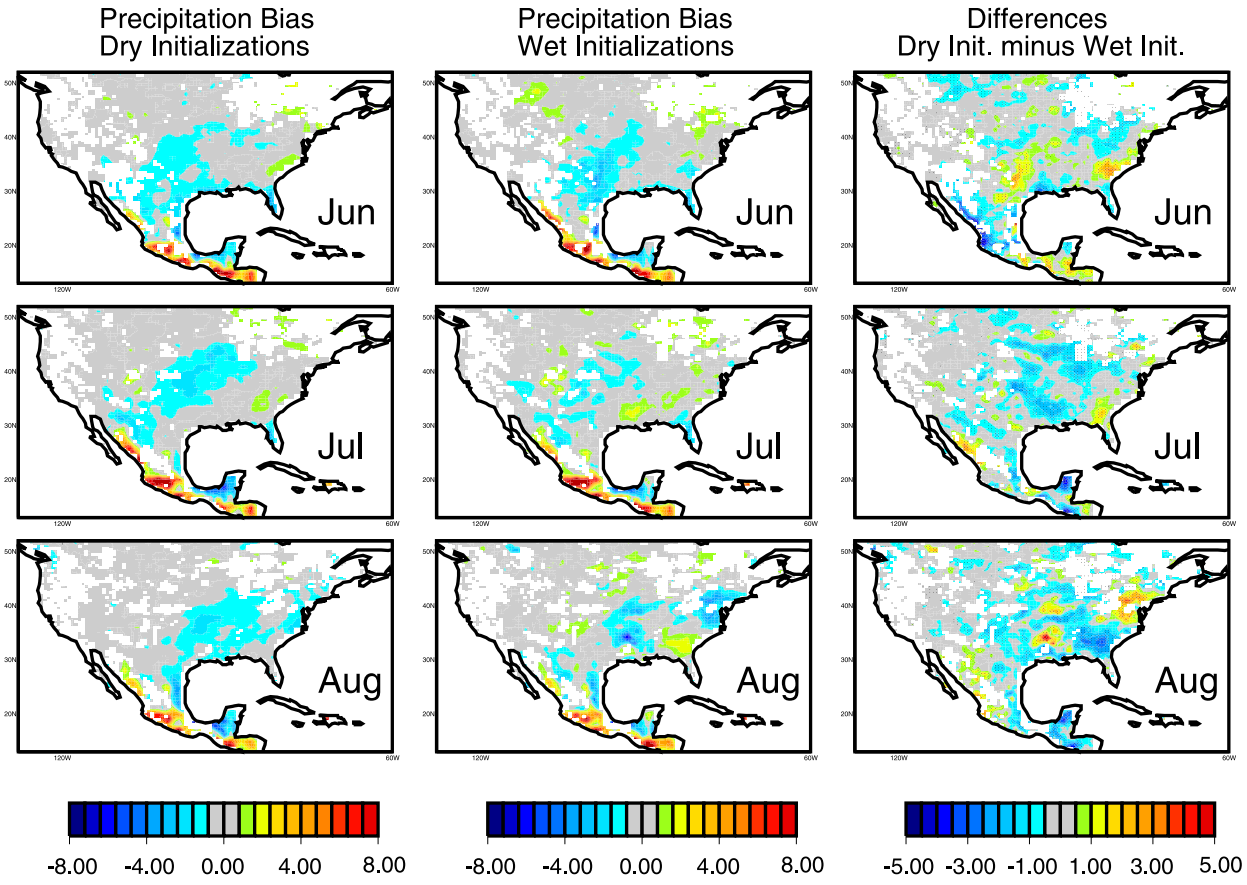

FIG. 7. (left) Precipitation forecast bias (mm day^{-1}) at the subseasonal lead for those forecasts with soil moisture initial states in the dry regime. These are composite maps; the particular forecasts contributing to the results shown are grid cell specific. (center) As in the left column, but for those forecasts with soil moisture initial states in the wet regime. (right) Differences: column 1 minus column 2. Dots indicate differences that are statistically significant at the 95% confidence level.

The discussion accompanying Fig. 3 suggests that the T2M ubRMSE should be larger for dry initial states than for wet initial states, and the difference plot for each month does indeed show a preponderance of positive values. In the area shown, grid cells with positive differences outnumber those with negative differences by a ratio of about 2.5 to 1, which is significant at the 95% level (section 2e). The spatially averaged difference for each month is roughly 0.25 K, which is significantly different from zero at the 99% level. The results thus lend support to the idea that the proposed ubRMSE-focused mechanism is indeed at play here.

The results, however, cannot be called conclusive; the differences in Fig. 8 are not as overwhelmingly positive as those seen in Fig. 6 for the bias-focused mechanism, and the individual locations (indicated by small dots) that are significantly different from zero at the 95% level are fewer in number. The differences are, in fact, generally much smaller than those seen for the bias-focused mechanism, as indicated by the different ranges for the color bars in Figs. 6 and 8. The smaller differences suggest that the impact of the ubRMSE-focused mechanism, if that mechanism is indeed operating, is more subtle than that of the bias-focused mechanism in this forecast system. A more extensive forecast suite, with a greater number of ensemble members per forecast, might be needed to demonstrate more conclusively this more subtle impact.

Central to the operation of the ubRMSE-focused mechanism is the idea that forecasts initialized dry should have a larger intraensemble spread in their subseasonal ET fluxes, to reflect the larger uncertainty of the actual ET. The availability of forecast output from the individual ensemble members thereby allows for an additional test of the mechanism—we can plot the differences (dry initial states minus wet initial states) in the average forecast ensemble spread in ET. Unfortunately, each individual forecast in our forecast suite consists of only four ensemble members, making the calculation of a standard deviation (or range) for that individual forecast highly inaccurate. Still, computing the intraensemble variance for each of the relevant forecasts in a given month and taking the square root of the average of the variances should provide an estimate of the standard deviation with first-order accuracy.

Figure 9 shows the differences (dry initial states minus wet initial states) in the average intraensemble ET standard deviation for each month over CONUS. The east–west gradient in the differences presumably reflects the fact that ET rates

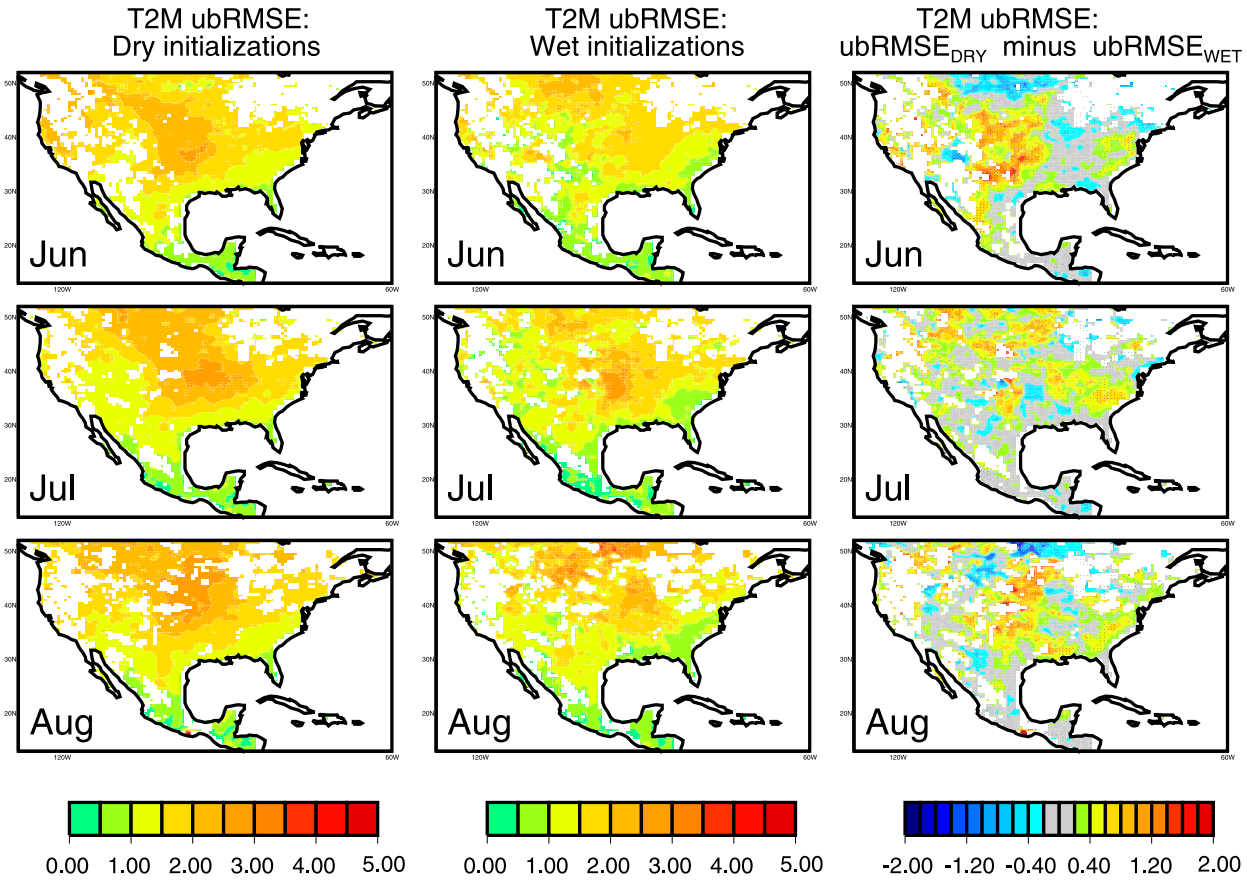

FIG. 8. (left) Average T2M forecast ubRMSE (K) generated by those forecasts with initial soil moisture in the dry regime (see Fig. 1). These are composite maps; the particular forecasts contributing to the results shown are grid cell specific. (center) As in the left column, but for those forecasts with initial soil moisture in the wet regime. (right) Differences: column 1 minus column 2. Dots indicate differences that are statistically significant at the 95% confidence level.

(and thus their variations) are higher in the east; while smaller, the differences in the west can still be significant, as indicated by the distribution of dots in the figure. The key message from Fig. 9 is that the differences across CONUS are almost entirely positive (with almost all of these differences being significant at the 95% level), supporting the idea that the proposed mechanism can indeed operate. Still, the positive differences in Fig. 9 do not translate into correspondingly consistent positive differences in the third column of Fig. 8. As suggested above, the ubRMSE-focused mechanism may have a relatively subtle impact on T2M random error, which would cloud the comparison—especially given that ET is only one of the factors controlling this error. Again, while the results in Figs. 8 and 9 generally support the idea that the ubRMSE-focused mechanism is in operation, a more extensive forecast suite with this forecast system or parallel analyses with other forecast systems (particularly systems with a larger ensemble size) may be needed to demonstrate this conclusively.

c. *Total error*

As indicated in Eq. (1), the forecast model bias and ubRMSE combine to produce the RMSE, a measure of the total forecast error. Figure 10 shows the T2M forecast RMSE for dry initial states minus that for wet initial states. Unlike bias, RMSE is a positive definite quantity, and thus the RMSE differences are readily interpreted in terms of forecast skill—positive (negative) differences imply that the dry initial state subset produces larger (smaller) total forecast errors.

For the central CONUS region, dry initial states clearly lead to higher total errors, and across CONUS, negative RMSE differences are small and infrequent. The differences, where they appear, are generally statistically significant at the 95% level. As noted above, comparison of the third columns in Figs. 6 and 8 indicate that for this forecast system, most of the RMSE differences in CONUS stem from bias differences.

4. Discussion

The mechanisms in Figs. 2 and 3 suggest that, before a forecast is even performed, the initial state of the soil can tell us something about the forecast's expected accuracy. In general, through these mechanisms, dry initial states can be expected to promote a greater random error and, in the presence of precipitation bias, a greater bias error (though see further

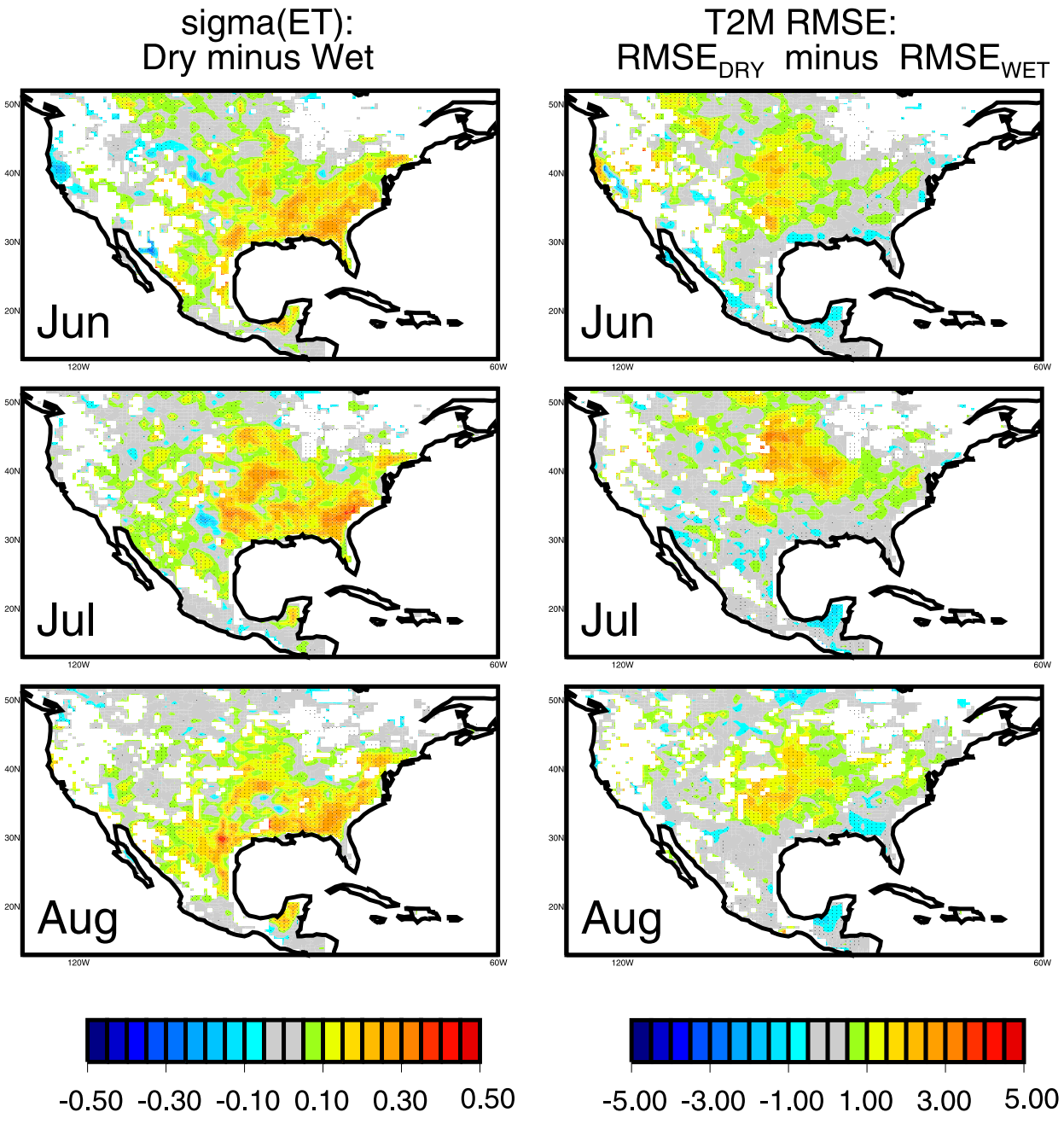

FIG. 9. Difference (dry initial states minus wet initial states) in the average intraensemble spread of forecasted ET (as represented by the standard deviation; mm day^{-1}). Dots indicate differences that are statistically significant at the 95% confidence level. These are composite maps; the particular forecasts contributing to the results shown are grid cell specific.

FIG. 10. Forecasted T2M RMSE differences (K): RMSE for those forecasts initialized in the dry regime minus that for forecasts initialized in the wet regime. Dots indicate differences that are statistically significant at the 95% confidence level. These are composite maps; the particular forecasts contributing to the results shown are grid cell-specific.

discussion below regarding the potential compensation of errors). Consideration of the mechanisms can thus contribute to the goal of identifying forecasts of opportunity—a cornerstone of our hopes for advancing subseasonal forecast skill (Mariotti et al. 2020). A wet soil moisture initial state may, under certain conditions, be an indicator of such a forecast of opportunity.

As an aside, we note that such ideas are, at first glance, consistent with the study of Huang and van den Dool (1993), who found higher antecedent rainfall rates to be associated with higher regression-based T2M forecast skill levels. This consistency, however, appears to be coincidental—Huang and van den Dool (1993) attributed the source of their wet-versus-dry

forecast skill differences to differences in the climatological periods used to generate their regressions, a factor not relevant in the present study.

Consideration of the mechanisms in Figs. 2 and 3 can be used for more than determining forecasts of opportunity. The idea that Earth system models are biased and that forecasts produced with such models can be made more realistic through bias correction is hardly new (e.g., Chang et al. 2019). The bias corrections applied, however, need not be constant; in principle they can be state-dependent (e.g., DelSole and Hou 1999; Danforth et al. 2007) or simply conditioned (as suggested by our results) on the initial state of the forecast system. Through consideration of the mechanism in Fig. 2, and guided by plots such as those in Fig. 6, forecasters could apply larger bias corrections to the T2M forecasts obtained under dry initializations, presumably leading to more accurate forecasts overall. The error that remains, the ubRMSE, would itself have some dependence on the initial land state (Fig. 8), being somewhat larger for drier initialization; through consideration of the ubRMSE-focused mechanism (Fig. 3), forecasters could assign higher quantitative uncertainty levels to these particular forecasts.

Given such potential applications, it is important to emphasize that our evaluation of the mechanisms over the CONUS region is not meant to imply that they are relevant only there. A more extensive forecast suite (covering a greater number of years) would have expanded somewhat the areas shaded red in Fig. 5a, and if our forecast system had not relied almost exclusively on the geographically limited rain gauge network for accuracy in the soil moisture initialization, our evaluations could have been more global in scale. Simply put, the mechanisms we propose presumably operate in many parts of the world. New datasets and approaches are expected, in the years to come, to improve considerably the initialization of soil moisture state in forecast systems, particularly in areas without good rain gauge coverage. Satellite-based rainfall estimates currently cover the globe at high (e.g., $0.1° \times 0.1°$) resolution (Huffman et al. 2020) and thus can help make up for rain gauge deficiencies. Remotely sensed measurements at L-band (Entekhabi et al. 2010b; Kerr et al. 2010) provide unprecedented global information on soil moisture conditions that can be assimilated into preforecast analysis systems (e.g., Carrera et al. 2019). With more accurate estimates of initial soil moisture state obtained through such techniques, the potential for the identification of forecasts of opportunity and the application of bias correction procedures need not be limited to areas with dense rain gauge networks.

All this being said, the Europe region outlined in Fig. 5c shows promise for some supplemental analyses with our current set of forecasts. While not as suitable as CONUS—according to a careful study of the panels in Fig. 5, only a region centered on the Balkan Peninsula satisfies all our analysis criteria—the rain gauge density in Europe is particularly high, and the CPC temperature validation data there are also of high quality. As background for some additional points of discussion, Fig. 11 shows relevant results over Europe for July. The precipitation bias (Fig. 11b) is negative in many parts of the region, particularly in the Balkan Peninsula, and these places are indeed where, at least for the most part, the difference in T2M forecast bias (dry initial states minus wet initial states) is positive (Fig. 11a). The mechanism in Fig. 2 thus does appear to be operating in Europe.

The operation in Europe of the ubRMSE-focused mechanism in Fig. 3, however, is not as clear. While the ubRMSE differences (dry initial states minus wet initial states) for July are generally positive as expected (Fig. 11c), the corresponding differences in the intraensemble standard deviation of ET forecasts (Fig. 11d) are much smaller than those in Fig. 9, being often close to zero. Errors in the estimation of the standard deviations from small forecast ensembles may explain the inconsistencies between Figs. 11c and 11d. Vagaries in how the single realization of nature manifests itself may also explain these inconsistencies; even with a large uncertainty represented by a large ensemble spread, nature may choose to produce a value close to the center of that range and thus close to the unbiased ensemble mean forecast. Thus, if the sample size underlying the ubRMSE calculation is small (as it often is), such vagaries could lead to a mismatch between ubRMSE and ensemble spread. We emphasize again, though, that Europe is not an optimal place to search for evidence of the mechanism. The inconsistency may simply stem from the low Corr(W, ET) values there (for this model) discussed in conjunction with Fig. 5b. Two findings support this idea: (i) the differences in Fig. 11d are generally positive as expected in the Balkan Peninsula, where all our analysis criteria are met, and (ii) the June version of Fig. 11d (not shown) exhibits a greater number of negative standard deviation differences, with June also having particularly low Corr(W, ET) values in Europe (also not shown).

The map of total RMSE differences in Europe (dry initial states minus wet initial states) shown in Fig. 11e illustrates another important discussion point. Taken at face value, both the bias-focused and ubRMSE-focused mechanisms seem to imply higher overall error for dry initial states. The total RMSE for T2M along the northern edge of Europe, however, is seen in Fig. 11e to be slightly lower for dry initial states. An analysis of overall climate biases in Europe reveals the reason: modeled temperatures in the northern half of Europe are already biased low in this forecast system (Fig. 11f) for reasons apparently unrelated to soil moisture – reasons possibly associated with the system's radiation balance. The small but negative precipitation bias seen in the northern half of Europe (Fig. 11b) presumably does encourage, through the mechanism in Fig. 2, a positive increase in the T2M bias under dry initial states (Fig. 11a), but this appears to mitigate, at least partially, the existing background cold T2M bias. In other words, in the northern half of Europe, it is a compensation of errors that allows the forecast T2M bias (and thus the total RMSE) to be smaller in magnitude for dry initial states than for wet initial states. The negative differences in Fig. 11e are thus not inconsistent with our proposed mechanisms. Rather, the Europe example highlights an important point: in some areas, it is critical to consider the soil moisture impacts on bias in conjunction with other sources of bias in the forecast system.

5. Summary

Our results show that the soil moisture regime (dry versus wet) characterizing a subseasonal forecast's initialization has a

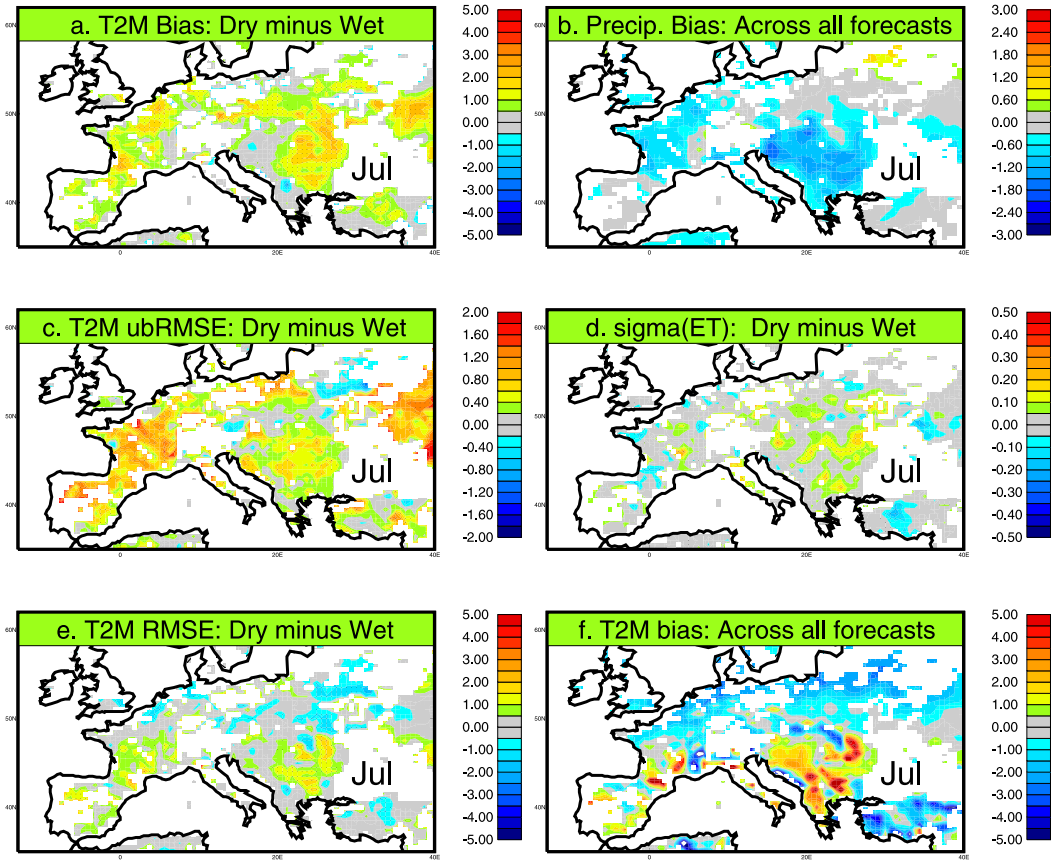

FIG. 11. Relevant difference plots for the Europe region. (a) Differences (dry initial states minus wet initial states) in the T2M forecast bias (K). Dots indicate differences that are statistically significant at the 95% confidence level. (b) Precipitation bias (mm day^{-1}) across all forecasts. (c) Differences (dry initial states minus wet initial states) in the T2M forecast ubRMSE (K). (d) Differences (dry initial states minus wet initial states) between the intraensemble spread in forecasted ET (mm day^{-1}). (e) Differences (dry initial states minus wet initial states) in the total T2M forecast RMSE. (f) Temperature bias (K) across all forecasts. All "Dry minus Wet" maps are composites; the particular forecasts contributing to the results shown are grid cell specific.

significant impact on T2M forecast bias. While not as conclusive, our results also lend support to the idea that soil moisture regime affects T2M random forecast error, as represented by the ubRMSE. Soil moisture regime thus has a significant impact on total T2M forecast error (Fig. 10), an impact that can be readily understood (through Figs. 2 and 3) in the context of the nonlinear relationship between soil moisture and ET efficiency. If one has already quantified the overall forecast biases (both temperature and precipitation) of a given modeling system through the analysis of historical forecasts, the mechanisms in Figs. 2 and 3 should provide some guidance regarding the distinction between T2M forecast error under dry and wet initial states. An initial soil moisture known at the outset to encourage lower forecast error can effectively point to a "forecast of opportunity"—a forecast that can be considered more accurate. Presumably, forecasters could use the mechanisms to derive bias corrections and uncertainty estimates conditioned on initial soil moisture state.

Naturally, progress in characterizing forecast error is critical for providing users with useful, ancillary information on forecast fidelity along with the forecasts themselves. It also provides some of the basic scientific underpinnings necessary to understand subseasonal predictability in the Earth system. The present study addresses an important and (to our knowledge) heretofore unexplored piece of this error characterization.

Acknowledgments. This work was supported by the NASA MAP (NNG17HP01C and WBS802678.02.17.01.33) and NOAA MAPP (NA14OAR4310221) programs. The forecasts and the MERRA-2 reanalysis were produced by the Global Modeling and Assimilation Office of NASA/GSFC, which is funded by the NASA MAP program.

Data availability statement. MERRA-2 data are disseminated through the Goddard Earth Science Data and Information Services Center (GES DISC). CPC Global Temperature data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, United States, from their website at https://www.esrl.noaa.gov/psd/. Forecast output may be obtained through the SubX project from the Data Library of the International

Research Institute for Climate and Society, Columbia University (http://iridl.ldeo.columbia.edu/SOURCES/.Models/.SubX/).

REFERENCES

Akbar, R., and Coauthors, 2018: Estimation of landscape soil water losses from satellite observations of soil moisture. *J. Hydrometeor.*, **19**, 871–889, https://doi.org/10.1175/JHM-D-17-0200.1.

Carrera, M. L., B. Bilodeau, S. Belair, M. Abrahamowicz, A. Russell, and X. Wang, 2019: Assimilation of passive L-band microwave brightness temperatures in the Canadian land data assimilation system: Impacts on short-range warm season numerical weather prediction. *J. Hydrometeor.*, **20**, 1053–1079, https://doi.org/10.1175/JHM-D-18-0133.1.

Chang, Y., S. D. Schubert, R. D. Koster, A. M. Molod, and H. Wang, 2019: Tendency bias correction in coupled and uncoupled global climate models with a focus on impacts over North America. *J. Climate*, **32**, 639–661, https://doi.org/10.1175/JCLI-D-18-0598.1.

Chen, M., W. Shi, P. Xie, V. B. S. Silva, V. E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110, https://doi.org/10.1029/2007JD009132.

Danforth, C. M., E. Kalnay, and T. Miyoshi, 2007: Estimating and correcting global weather model error. *Mon. Wea. Rev.*, **135**, 281–299, https://doi.org/10.1175/MWR3289.1.

DelSole, T., and A. Y. Hou, 1999: Empirical correction of a dynamical model. Part I: Fundamental issues. *Mon. Wea. Rev.*, **127**, 2533–2545, https://doi.org/10.1175/1520-0493(1999)127<2533:ECOADM>2.0.CO;2.

Dirmeyer, P. A., and S. Halder, 2017: Application of the land–atmosphere coupling paradigm to the operational Coupled Forecast System, Version 2 (CFSv2). *J. Hydrometeor.*, **18**, 85–108, https://doi.org/10.1175/JHM-D-16-0064.1.

——, R. D. Koster, and Z. Guo, 2006: Do global models properly represent the feedback between the land and atmosphere? *J. Hydrometeor.*, **7**, 1177–1198, https://doi.org/10.1175/JHM532.1.

Eagleson, P. S., 1978: Climate, soil and vegetation, 4. The expected value of annual evapotranspiration. *Water Resour. Res.*, **14**, 731–739, https://doi.org/10.1029/WR014i005p00731.

Entekhabi, D., R. H. Reichle, R. D. Koster, and W. Crow, 2010a: Performance metrics for soil moisture retrievals and application requirements. *J. Hydrometeor.*, **11**, 832–840, https://doi.org/10.1175/2010JHM1223.1.

Entekhabi, D. E., and Coauthors, 2010b: The Soil Moisture Active Passive (SMAP) mission. *Proc. IEEE*, **98**, 704–716, https://doi.org/10.1109/JPROC.2010.2043918.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

GMAO, 2015a: tavg1_2d_lnd_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), accessed 27 September 2019, https://doi.org/10.5067/RKPHT8KC1Y1T.

——, 2015b: tavg1_2d_flx_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), accessed 27 September 2019, https://doi.org/10.5067/7MCPBJ41Y0K6.

——, 2015c: tavg1_2d_rad_Nx hourly collection. Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), accessed 27 September 2019, https://doi.org/10.5067/Q9QMY5PBNV1T.

Hirsch, A. L., and Coauthors, 2014: Impact of land surface initialization approach on subseasonal forecast skill: A regional analysis in the Southern Hemisphere. *J. Hydrometeor.*, **15**, 300–319, https://doi.org/10.1175/JHM-D-13-05.1.

Huang, J., and H. M. van den Dool, 1993: Monthly precipitation-temperature relations and temperature prediction over the United States. *J. Climate*, **6**, 1111–1132, https://doi.org/10.1175/1520-0442(1993)006<1111:MPTRAT>2.0.CO;2.

Huffman, G. J., and Coauthors, 2020: Integrated multi-satellitE retrievals for the Global Precipitation Measurement (GPM) mission (IMERG). *Satellite Precipitation Measurement*, V. Levizzani et al., Eds., Advances in Global Change Research, Vol. 67, Springer, 343–353.

Kerr, Y. H., and Coauthors, 2010: The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proc. IEEE*, **98**, 666–687, https://doi.org/10.1109/JPROC.2010.2043032.

Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum, 2017: So, how much of the Earth's surface is covered by rain gauges? *Bull. Amer. Meteor. Soc.*, **98**, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1.

Kim, H., F. Vitart, and D. E. Waliser, 2018: Prediction of the Madden–Julian oscillation: A review. *J. Climate*, **31**, 9425–9443, https://doi.org/10.1175/JCLI-D-18-0210.1.

Koster, R. D., and M. J. Suarez, 2001: Soil moisture memory in climate models. *J. Hydrometeor.*, **2**, 558–570, https://doi.org/10.1175/1525-7541(2001)002<0558:SMMICM>2.0.CO;2.

——, and ——, 2003: Impact of land surface initialization on seasonal precipitation and temperature prediction. *J. Hydrometeor.*, **4**, 408–423, https://doi.org/10.1175/1525-7541(2003)4<408:IOLSIO>2.0.CO;2.

——, ——, A. Ducharne, M. Stieglitz, and P. Kumar, 2000a: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.*, **105**, 24 809–24 822, https://doi.org/10.1029/2000JD900327.

——, ——, and M. Heiser, 2000b: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeor.*, **1**, 26–46, https://doi.org/10.1175/1525-7541(2000)001<0026:VAPOPA>2.0.CO;2.

——, and Coauthors, 2011: The second phase of the Global Land-Atmosphere Coupling Experiment: Soil moisture contributions to subseasonal forecast skill. *J. Hydrometeor.*, **12**, 805–822, https://doi.org/10.1175/2011JHM1365.1.

——, Y. Chang, H. Wang, and S. D. Schubert, 2016: Impacts of local soil moisture anomalies on the atmospheric circulation and on remote surface meteorological fields during boreal summer: A comprehensive analysis over North America. *J. Climate*, **29**, 7345–7364, https://doi.org/10.1175/JCLI-D-16-0192.1.

——, S. D. Schubert, H. Wang, S. P. Mahanama, and A. M. DeAngelis, 2019: Flash drought as captured by reanalysis data: Disentangling the contributions of precipitation deficit and excess evapotranspiration. *J. Hydrometeor.*, **20**, 1241–1258, https://doi.org/10.1175/JHM-D-18-0242.1.

——, ——, A. M. DeAngelis, A. M. Molod, and S. P. P. Mahanama, 2020: Using a simple water balance framework to quantify the impact of soil moisture initialization on subseasonal evapotranspiration and air temperature forecasts. *J. Hydrometeor.*, **21**, 1705–1722, https://doi.org/10.1175/JHM-D-20-0007.1.

Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227**, 3515–3539, https://doi.org/10.1016/j.jcp.2007.02.014.

Lin, Y., and Coauthors, 2017: Causes of model dry and warm bias over central U.S. and impact on climate projections. *Nat. Commun.*, **8**, 881, https://doi.org/10.1038/s41467-017-01040-2.

Lucchesi, R., 2018: File specification for GEOS-5 FP (Forward Processing). GMAO Office Note 4, version 1.2, 61 pp., https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi1203.pdf.

Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bull. Amer. Meteor. Soc.*, **101**, E608–E625, https://doi.org/10.1175/BAMS-D-18-0326.1.

McCarty, W., L. Coy, R. Gelaro, A. Huang, D. Merkova, E. B. Smith, M. Sienkiewicz, and K. Wargan, 2016: MERRA-2 input observations: Summary and assessment. NASA/TM-2016-104606, Vol. 46, 61 pp., https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20160014544.pdf.

Molod, A. M., L. Takacs, M. Suárez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339–1356, https://doi.org/10.5194/gmd-8-1339-2015.

——, and Coauthors, 2020: GEOS-S2S version 2: The GMAO high-resolution coupled model and assimilation system for seasonal prediction. *J. Geophys. Res. Atmos.*, **125**, e2019JD031767, https://doi.org/10.1029/2019JD031767.

Moorthi, S., and M. J. Suarez, 1992: Relaxed Arakawa–Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.*, **120**, 978–1002, https://doi.org/10.1175/1520-0493(1992)120<0978:RASAPO>2.0.CO;2.

Pegion, K., and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, https://doi.org/10.1175/BAMS-D-18-0270.1.

Randles, C. A., and Coauthors, 2016: The MERRA-2 aerosol assimilation. NASA Tech. Rep. Series on Global Modeling and Data Assimilation. NASA/TM-2016-104606, Vol. 45, 153 pp., https://gmao.gsfc.nasa.gov/pubs/docs/Randles887.pdf.

Reichle, R. H., and Coauthors, 2017: Land surface precipitation in MERRA-2. *J. Climate*, **30**, 1643–1664, https://doi.org/10.1175/JCLI-D-16-0570.1.

Robertson, A. W., A. Kumar, M. Peña, and F. Vitart, 2015: Improving and promoting subseasonal to seasonal prediction. *Bull. Amer. Meteor. Soc.*, **96**, ES49–ES53, https://doi.org/10.1175/BAMS-D-14-00139.1.

Ruiz-Barradas, A., and S. Nigam, 2006: Great Plains hydroclimate variability: The view from North American Regional Reanalysis. *J. Climate*, **19**, 3004–3010, https://doi.org/10.1175/JCLI3768.1.

Salvucci, G. D., 2001: Estimating the moisture dependence of root zone water loss using conditionally averaged precipitation. *Water Resour. Res.*, **37**, 1357–1365, https://doi.org/10.1029/2000WR900336.

Scaife, A. A., and Coauthors, 2016: Seasonal winter forecasts and the stratosphere. *Atmos. Sci. Lett.*, **17**, 51–56, https://doi.org/10.1002/asl.598.

Seneviratne, S., and Coauthors, 2006: Soil moisture memory in AGCM simulations: Analysis of Global Land-Atmosphere Coupling Experiment (GLACE) data. *J. Hydrometeor.*, **7**, 1090–1112, https://doi.org/10.1175/JHM533.1.

——, and Coauthors, 2010: Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004.

Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science*, **282**, 728–731, https://doi.org/10.1126/science.282.5389.728.

van den Hurk, B., and Coauthors, 2012: Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dyn.*, **38**, 349–362, https://doi.org/10.1007/s00382-010-0956-2.

Vitart, F., and Coauthors, 2017: The Subseasonal to Seasonal (S2S) Prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, https://doi.org/10.1175/BAMS-D-16-0017.1.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558, https://doi.org/10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

——, M. Chen, S. Yang, A. Yatagai, T. Hayasaka, Y. Fukushima, and C. Liu, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626, https://doi.org/10.1175/JHM583.1.

Zhang, F., and Coauthors, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091, https://doi.org/10.1175/JAS-D-18-0269.1.